

EXHIBIT B

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

----- X
THE NEW YORK TIMES COMPANY, : No. 1:23-cv-11195-SHS
Plaintiff, :
: v.
MICROSOFT CORPORATION, OPENAI, :
INC., OPENAI LP, OPENAI GP, LLC, :
OPENAI, LLC, OPENAI OPCO LLC, :
OPENAI GLOBAL LLC, OAI :
CORPORATION, LLC, and OPENAI X
HOLDINGS, LLC,
Defendants.

**OPENAI DEFENDANTS' RESPONSES AND OBJECTIONS TO PLAINTIFF'S FIRST
SET OF REQUESTS FOR PRODUCTION OF DOCUMENTS**

Pursuant to Rules 26 and 34 of the Federal Rules of Civil Procedure and the Local Rules of the United States District Courts for the Southern and Eastern Districts of New York (“Local Rules”), Defendants OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OpCo, LLC, OpenAI Global, LLC, OAI Corporation, LLC and OpenAI Holdings, LLC (collectively, “OpenAI”) hereby object and respond to Plaintiff’s First Set of Requests for Production of Documents (“Requests”). To the extent that OpenAI agrees to produce documents in response to these Requests, each entity is agreeing to produce only its own documents, to the extent those documents can be located after a reasonable search. Furthermore, an agreement by OpenAI to search for documents does not mean that each entity has documents in its possession, custody, or control.

INTRODUCTORY RESPONSE

OpenAI responds to the Requests on the basis of the best information available to it at the

time the responsive information was gathered, within the limits of time, and subject to the objections described below. OpenAI responds to the Requests as it interprets and understands each Request set forth herein. If Plaintiff subsequently asserts an interpretation of any of the Requests that differs from OpenAI's understanding, OpenAI reserves the right to supplement its objections and/or responses.

OpenAI's willingness to respond to any particular Request does not constitute an admission that OpenAI agrees with any characterization, definition, or assumption contained in the Request or an assumption or an acknowledgement by OpenAI that the Request is proper, that the information sought is within the proper bounds of discovery or that demands for similar information will be treated in similar fashion. Furthermore, a statement that responsive documents will be produced in response to a particular Request does not mean that OpenAI knows any such document exists or is in its possession, custody, or control.

OpenAI's responses to the Requests may contain, provide, or refer to information that will be subject to a protective order to be entered in this matter and should therefore be treated accordingly.

OpenAI has not yet completed its discovery relating to this case and its investigation of the facts is ongoing. OpenAI anticipates that additional information responsive to the Requests may be obtained as discovery proceeds. OpenAI's responses to the Requests are therefore made without prejudice to OpenAI's right to amend, correct or supplement its responses to the Requests.

GENERAL OBJECTIONS

1. OpenAI objects to the Requests, including the Definitions and Instructions contained therein, to the extent that they:

- a. call for the discovery of information protected by the attorney-client privilege, the attorney work-product immunity and/or any other applicable privilege or immunity under federal, state or local law. The inadvertent production of any information that is privileged or otherwise protected from discovery shall not constitute a waiver of any such privilege or protection;
- b. are vague, ambiguous, overly broad, unduly burdensome, unreasonably cumulative, overlapping, or duplicative;
- c. require disclosure of information that is neither relevant to the claims or defenses of any party in this action, nor proportional to the needs of the case;
- d. purport to require the disclosure of information already in the possession of Plaintiff, available from public sources, as accessible to Plaintiff as to OpenAI, or obtainable from another source that is more convenient, less burdensome or less expensive. OpenAI provides these responses with the understanding that Plaintiff already has access to such sources, including all materials on OpenAI's website;
- e. attempt to impose burdens that exceed or are different from the requirements of the Federal Rules of Civil Procedure and/or the applicable Local Rules, and relevant case law. OpenAI will comply with its obligations under the Federal and Local Rules, but will not comply with Instructions that are inconsistent with, or impose obligations beyond, applicable rules. OpenAI stands ready to meet and confer to discuss the best and most efficient way to conduct discovery bilaterally;
- f. call for information concerning documents that are no longer in existence or reasonably accessible, or documents not in OpenAI's possession, custody, or control. OpenAI will only provide relevant, non-privileged information that is within OpenAI's

present possession, custody, or control and available after a reasonable investigation;

g. imprecisely specify the information sought. Where imprecise terms are used, OpenAI will, therefore, provide only information that is responsive to the Request based on OpenAI's interpretation of the Request;

h. seek information concerning "all documents and things," or the like on the basis that providing a response to such Requests would be unduly burdensome;

i. are not limited to a reasonable time period. OpenAI will produce documents from a reasonable time period as it relates to the case;

j. purport to impose a duty on OpenAI to undertake a search for information beyond a reasonable search of its files as it relates to this case and where information responsive to the Request would reasonably be expected to be stored in the ordinary course of business;

k. seek discovery in violation of any applicable law;

l. are premature and/or the disclosure of expert discovery, opinions or analysis;

m. seek information that is the confidential and proprietary information of a third party, the joint confidential and proprietary information of OpenAI and a third party, or the subject of a non-disclosure/confidentiality agreement between OpenAI and a third party. OpenAI will produce third-party confidential documents as set forth herein, subject to the terms of the protective order to be entered by the Court; or seek information that is more efficiently or appropriately obtained through some other form of discovery.

2. Where the Request includes words and concepts involving a legal conclusion, any response is not an admission that such legal conclusions apply.

3. OpenAI reserves all objections to the competency, relevance, materiality, or admissibility at trial or any other proceeding of any information provided or document identified in response to the Requests. The identification of any document or the provision of any information does not constitute an admission that such document or information exists or is relevant to the pending litigation.

4. OpenAI objects to Plaintiff's definitions of "Defendants," "OpenAI," "You" and "Your" to the extent they purport to include each of the OpenAI Defendants' "subsidiaries, affiliates, divisions, predecessors in interest, successors, or assignees, and their respective officers, directors, employees, consultants, representatives, and agents" as overly broad, unduly burdensome and disproportionate to the needs of this case and beyond the scope of permissible discovery under the Federal Rules of Civil Procedure and the Local Rules. OpenAI interprets and defines OpenAI to include only the named Defendants OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OpCo, LLC, OpenAI Global, LLC, OAI Corporation, LLC and OpenAI Holdings, LLC.

5. OpenAI objects to Plaintiff's definition of "AI Model(s)," "Generative AI Model(s)," "Text Generation AI Model(s)," and "Generative AI Product(s) and Service(s)" as overly broad, unduly burdensome, and not proportional to the needs of the case. Whereas these definitions encompass generative AI and AI models generally, Plaintiff's allegations against OpenAI implicate only the models used for ChatGPT and are more limited in scope. OpenAI interprets and defines "AI Model(s)", "Generative AI Model(s)," and "Text Generation AI Model(s)" to include only the models used for ChatGPT. OpenAI interprets and defines "Generative AI Product(s) and Service(s)" as ChatGPT.

6. OpenAI objects to Plaintiff's definition of "Journalism" as overly broad, unduly

burdensome, and not proportional to the needs of the case. OpenAI interprets and defines “Journalism” as “the collection and editing of news for presentation through the media.” “journalism.” *Merriam-Webster.com*. 2024. <https://www.merriam-webster.com/dictionary/journalism>.

7. OpenAI objects to the requested production date as unreasonably burdensome.
8. OpenAI will produce documents at a reasonable time consistent with the case schedule and an entry of a protective order in this matter.
9. OpenAI objects to each Request to the extent that it seeks trade secrets or confidential or proprietary information of OpenAI. OpenAI will produce confidential documents as set forth herein, subject to and after the entry of a protective order in this matter.
10. OpenAI objects to the production of electronically stored information (“ESI”) before an applicable ESI protocol has been entered. OpenAI further objects to Plaintiff’s instructions regarding production of ESI as misrepresenting any agreements the parties have entered into regarding ESI. OpenAI will produce electronically stored information as set forth herein, subject to the terms of an ESI protocol to be agreed upon by the parties.
11. OpenAI objects to requests to the extent they purport to require the OpenAI defendants to each produce their own copy of a responsive document as duplicative, unreasonably burdensome and not proportional to the needs of this case.
12. OpenAI objects to the instruction that purports to require OpenAI to produce documents not otherwise responsive to these Requests if the documents relate to the documents called for by these Requests as seeking irrelevant information, unreasonably burdensome, and not proportional to the needs of this case.
13. OpenAI objects to the instruction that purports to require OpenAI to identify

documents that have been lost, discarded or destroyed, the date of disposal, the manner of disposal, the reason for disposal, the person authorizing disposal, and the person disposing of the document. OpenAI will not identify such documents, as the requests to do so is unreasonably burdensome and not proportional to the needs of this case.

Without waiving the foregoing General Objections but in express reliance thereon, OpenAI incorporates the foregoing objections into the responses below and responds to the individually numbered Requests as follows:

SPECIFIC OBJECTIONS AND RESPONSES

REQUEST FOR PRODUCTION NO. 1:

Documents concerning the use of Times Content for training or fine tuning Your Text Generation AI Models, including, for every Text Generation AI Model that uses Times Content: i) the amount and prevalence of Times Content in the Training Datasets by token count, word count, article count, as a percentage of the Training Datasets, or by any other metric; ii) rules or guidance for choosing what is and is not included in the Training Datasets for the Text Generation AI Models; iii) use of Times Content to clean or curate the Training Datasets; iv) any weighting of particular Training Datasets for the Text Generation AI Models; and v) use of Times Content for any post-training activity, including fine tuning.

RESPONSE TO REQUEST FOR PRODUCTION NO. 1:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the request for “the amount and prevalence of Times Content in the Training Datasets by token word, word count, article count, as a percentage of the Training Datasets, or by any other

metrics” to the extent it imposes an obligation to create information and/or documents that do not already exist. OpenAI objects to the phrases “Times Content,” “use of Times Content,” “uses Times Content,” and “rules or guidance” as vague and ambiguous. OpenAI objects to this request as overly broad, unduly burdensome, disproportionate to the needs of the case to the extent this request seeks documents regarding models not used for ChatGPT. OpenAI further objects to this request as seeking information and/or documents that are not relevant to any claim or defense in this litigation to the extent this request seeks documents regarding models not used for ChatGPT. OpenAI objects to this request to the extent it seeks information protected by the attorney-client privilege, work-product doctrine, common-interest privilege, or any other applicable privilege or protection.

Subject to the foregoing General and Specific Objections, OpenAI will make available for inspection, pursuant to an inspection protocol to be negotiated between the parties, the pretraining data for models used for ChatGPT that it locates after a reasonable search.

REQUEST FOR PRODUCTION NO. 2:

Documents You have gathered to submit or have submitted to any legislative or executive agency, committee, or other governmental entity in the United States that concern or relate to the allegations in the Complaint, including any Documents related to Your intended or actual submission.

RESPONSE TO REQUEST FOR PRODUCTION NO. 2:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI further objects to the request for “Documents related to Your intended . . . submission,” as overly broad, unduly

burdensome, and disproportionate to the needs of this case. OpenAI objects to this request as seeking information and/or documents that are not relevant to any claim or defense in this litigation. OpenAI objects to this request to the extent it seeks information protected by the attorney-client privilege, work-product doctrine, common-interest privilege, or any other applicable privilege or protection. OpenAI objects to this request to the extent it seeks confidential or proprietary information. OpenAI objects to this request to the extent it seeks information protected from disclosure by the First Amendment.

Subject to the foregoing General and Specific Objections, OpenAI states that it will not produce documents in response to this request.

REQUEST FOR PRODUCTION NO. 3:

Documents concerning copyrighted material in the Text Generation AI Models, the Training Datasets, or any Generative AI Products and Services, including documents or communications regarding the appropriateness, value, importance, necessity, and/or concerns over or legal implications of the use of such materials.

RESPONSE TO REQUEST FOR PRODUCTION NO. 3:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the phrases “Documents concerning copyrighted material,” “appropriateness,” “value,” “importance,” “necessity,” “concerns over,” and “legal implications” as vague and ambiguous. OpenAI further objects to this request as seeking information and/or documents that are not relevant to any claim or defense in this litigation. OpenAI objects to this request to the extent it

seeks information protected by the attorney-client privilege, work-product doctrine, common-interest privilege, or any other applicable privilege or protection.

Subject to the foregoing General and Specific Objections, OpenAI will make available for inspection, pursuant to an inspection protocol to be negotiated between the parties, the pretraining data for models used for ChatGPT that it locates after a reasonable search.

REQUEST FOR PRODUCTION NO. 4:

Documents concerning claims of copyright and trademark infringement, including complaints regarding infringement, steps taken to respond to such claims, and steps taken to prevent infringement.

RESPONSE TO REQUEST FOR PRODUCTION NO. 4:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the phrases “claims,” “complaints,” “respond,” and “steps taken” as vague and ambiguous. OpenAI objects to this request to the extent it seeks information protected by the attorney-client privilege, work-product doctrine, common-interest privilege, or any other applicable privilege or protection.

Subject to the foregoing General and Specific Objections, OpenAI states that it will produce non-privileged documents in its possession, custody, or control, if any, that it locates pursuant to a reasonable and diligent search sufficient to show (a) OpenAI’s procedures governing formal, written complaints regarding the use of purportedly copyrighted and trademarked material, (b) the receipt of such complaints, if any, and (c) OpenAI’s responses, if any, to such complaints, after the entry of a protective order in this case.

REQUEST FOR PRODUCTION NO. 5:

Documents reflecting policies, procedures or practices concerning use of intellectual property in AI Models, including the policy OpenAI referred to as “Copyright Shield” in the November 6, 2023, blog post on Your website titled “New models and developer products announced at DevDay.”

RESPONSE TO REQUEST FOR PRODUCTION NO. 5:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the phrase “use of intellectual property” as vague and ambiguous. OpenAI further objects to the request for “Documents reflecting policies, procedures, or practice concerning use of intellectual property in AI Models,” as overly broad, unduly burdensome, and disproportionate to the needs of this case. OpenAI further objects to this request as seeking information and/or documents that are not relevant to any claim or defense in this litigation to the extent this request seeks “Documents . . . concerning use of intellectual property” beyond copyright and trademark. OpenAI objects to this request to the extent it seeks information protected by the attorney-client privilege, work-product doctrine, common-interest privilege, or any other applicable privilege or protection.

Subject to the foregoing General and Specific Objections, OpenAI states that it will produce non-privileged documents in its possession, custody, or control, if any, that it locates pursuant to a reasonable and diligent search sufficient to explain the “Copyright Shield” policy, after the entry of a protective order in this case.

REQUEST FOR PRODUCTION NO. 6:

Documents concerning alternatives to using copyrighted content to train the Text Generation AI Models without compensation including using licensing agreements or material in the public domain.

RESPONSE TO REQUEST FOR PRODUCTION NO. 6:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the phrase “alternatives” as vague and ambiguous. OpenAI further objects to this request as seeking information and/or documents that are not relevant to any claim or defense in this litigation. OpenAI objects to this request to the extent it seeks information protected by the attorney-client privilege, work-product doctrine, common-interest privilege, or any other applicable privilege or protection.

Subject to the foregoing General and Specific Objections, OpenAI states that it will produce any responsive, non-privileged documents in its possession, custody, or control, if any, that it locates pursuant to a reasonable and diligent search constituting agreements with third parties relating to training data that was used to train the models used for ChatGPT, after the entry of a protective order.

REQUEST FOR PRODUCTION NO. 7:

Documents concerning the impact of AI Models on Journalism and the past, current or future use of the Text Generation AI Models or any Generative AI Products and Services as a source of Journalism, substitute for Journalism, or product in competition with Journalism (including The Times’s Journalism).

RESPONSE TO REQUEST FOR PRODUCTION NO. 7:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to Plaintiffs' definition of "Journalism," as defined in the Requests as "the activity of writing or creating content for newspapers, magazines, news websites, mobile applications, television, podcasts, or any other publication and/or news outlet" as overly broad, unduly burdensome, and not proportional to the needs of the case. OpenAI objects to the phrases "impact of AI Models on Journalism," "source," "substitute," or "product in competition" as vague and ambiguous. OpenAI objects to the request it calls for a legal conclusion. OpenAI objects to this request to the extent it seeks information protected by the attorney-client privilege, work-product doctrine, common-interest privilege, or any other applicable privilege or protection.

Subject to the foregoing General and Specific Objections, OpenAI states that it will produce non-privileged documents in its possession, custody, or control, if any, that it locates pursuant to a reasonable and diligent search sufficient to show OpenAI's understanding of ChatGPT's effect, if any, on journalism, to the extent such documents exist, after the entry of a protective order in this case.

REQUEST FOR PRODUCTION NO. 8:

Documents sufficient to show to what extent and in what form Times Content resides in, is stored in, is copied by, or otherwise exists within the Training Datasets and/or the Text Generation AI Models.

RESPONSE TO REQUEST FOR PRODUCTION NO. 8:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the phrases “resides in,” “stored in,” “copied by” and “exists within” as vague and ambiguous.

Subject to the foregoing General and Specific Objections, OpenAI will make available for inspection, pursuant to an inspection protocol to be negotiated between the parties, the pretraining data for models used for ChatGPT that it locates after a reasonable search.

REQUEST FOR PRODUCTION NO. 9:

Documents sufficient to show and describe OpenAI’s corporate organizational structure, including those sufficient to show OpenAI’s board of directors, parent companies, subsidiaries, affiliates, departments, and divisions; and those sufficient to show the function each Defendant plays within Your corporate structure as a whole.

RESPONSE TO REQUEST FOR PRODUCTION NO. 9:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the request for documents regarding “OpenAI’s board of directors, parent companies, subsidiaries, affiliates, departments, and divisions” as overly broad, unduly burdensome, and not proportional to the needs of the case. OpenAI objects to the phrase “the function each Defendant plays” as vague and ambiguous. OpenAI objects to the extent this request seeks information and/or documents regarding entities that are not relevant to any claim or defense in this litigation.

Subject to the foregoing General and Specific Objections, OpenAI states that information about its Board of Directors and corporate structure is publicly available on its website, including at <https://openai.com/blog> and <https://openai.com/our-structure>. OpenAI further states that it will produce non-privileged documents in its possession, custody, or control, if any, that it locates pursuant to a reasonable and diligent search sufficient to show the corporate relationship between the different defendant OpenAI entities, after the entry of a protective order in this case. To the extent Plaintiff believes it is entitled to additional information, OpenAI is willing to meet and confer regarding an appropriately narrow scope for this request.

REQUEST FOR PRODUCTION NO. 10:

Documents sufficient to identify OpenAI's current and former officers, directors, managers, employees and consultants with knowledge of the Text Generation AI Models, business practices regarding the Text Generation AI Models, and commercialization and use of the Text Generation AI Models.

RESPONSE TO REQUEST FOR PRODUCTION NO. 10:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the request for documents regarding "OpenAI's current and former officers, directors, managers, employees, and consultants" who have knowledge of its "Text Generation AI Models, business practices regarding the Text Generation AI Models, and commercialization and use of the Text Generation AI Models" as overly broad, unduly burdensome, and not proportional to the needs of the case. OpenAI objects to the phrases "business practices regarding the Text Generation AI Models," "commercialization," and "use" and as vague and ambiguous. OpenAI objects to the

extent this request seeks information and/or documents regarding individuals that are not relevant to any claim or defense in this litigation and/or that are not proportional to the needs of this case.

Subject to the foregoing General and Specific Objections, OpenAI will produce nonprivileged documents in its possession, custody, or control, if any, that it locates pursuant to a reasonable and diligent search sufficient to show the roles of individuals who were primarily responsible for the design and operation of the models used for ChatGPT, after the entry of a protective order in this case.

REQUEST FOR PRODUCTION NO. 11:

Documents concerning OpenAI's transition to a for-profit company, as alleged in paragraphs 57-58 of the Complaint.

RESPONSE TO REQUEST FOR PRODUCTION NO. 11:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the phrase "transition to a for-profit company" as vague and ambiguous. OpenAI objects to the extent this request seeks information and/or documents regarding entities that are not relevant to any claim or defense in this litigation, and/or are not proportional to the needs of this case.

OpenAI objects to this request to the extent it seeks information protected by the attorney-client privilege, work-product doctrine, common-interest privilege, or any other applicable privilege or protection.

Subject to the foregoing General and Specific Objections, OpenAI states that it will produce non-privileged documents in its possession, custody, or control, if any, that it locates

pursuant to a reasonable and diligent search to show whether each OpenAI Defendant is a non-profit or a for-profit entity, including documents sufficient to show any changes to the entity's non-profit or for-profit status, to the extent any such changes have taken place.

REQUEST FOR PRODUCTION NO. 12:

Documents concerning Defendants' use of Retrieval Augmented Generation ("RAG") in Generative AI Products and Services, including: (i) use of Times Content for RAG; (ii) rules or guidance for preferring or avoiding certain web domains (including The Times's web domains or sites known to host content copied from other sources) when choosing what to include in a generative result; (iii) rules for whether and how the results from RAG search products are similar to or different from standard search results; (iv) and rules regarding paywalls.

RESPONSE TO REQUEST FOR PRODUCTION NO. 12:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the request for documents regarding "Defendants' use of Retrieval Augmented Generation" as overly broad, unduly burdensome, and not proportional to the needs of the case. OpenAI objects to the phrase "rules or guidance for preferring or avoiding certain web domains," "choosing what to include in a generative result," "standard search results," and "paywalls" as vague and ambiguous. OpenAI objects to this request to the extent it seeks information protected by the attorney-client privilege, work-product doctrine, common-interest privilege, or any other applicable privilege or protection.

Subject to the foregoing General and Specific Objections, OpenAI states that it will produce non-privileged documents in its possession, custody, or control, if any, that it locates

pursuant to a reasonable and diligent search sufficient to show how OpenAI uses RAG in connection with ChatGPT.

REQUEST FOR PRODUCTION NO. 13:

Documents concerning the features or capabilities of Retrieval Augmented Generation.

RESPONSE TO REQUEST FOR PRODUCTION NO. 13:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the request for documents regarding “features or capabilities of Retrieval Augmented Generation” as overly broad, unduly burdensome, and not proportional to the needs of the case. OpenAI objects to the terms “features” and “capabilities” as vague and ambiguous. OpenAI further objects to this request as seeking information and/or documents that are not relevant to any claim or defense in this litigation to the extent this request seeks “Documents concerning the features or capabilities of Retrieval Augmented Generation” beyond those related to the use of Retrieval Augmented Generation, if any, by ChatGPT. OpenAI objects to this request to the extent it seeks information protected by the attorney-client privilege, work-product doctrine, common-interest privilege, or any other applicable privilege or protection.

Subject to the foregoing General and Specific Objections, OpenAI states that it will produce non-privileged documents in its possession, custody, or control, if any, that it locates pursuant to a reasonable and diligent search sufficient to show OpenAI’s understanding of RAG.

REQUEST FOR PRODUCTION NO. 14:

A complete listing of every data sample in the Training Datasets used to train each of the Text Generation AI Models and the source from which each data sample was obtained.

RESPONSE TO REQUEST FOR PRODUCTION NO. 14:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the request to the extent it imposes an obligation to create information and/or documents that do not already exist. OpenAI objects to the phrases “data sample” and “source” as vague and ambiguous. OpenAI objects to this request as overly broad, unduly burdensome, disproportionate to the needs of the case to the extent this request seeks documents regarding models not used for ChatGPT. OpenAI further objects to this request as seeking information and/or documents that are not relevant to any claim or defense in this litigation to the extent this request seeks documents regarding models not used for ChatGPT.

Subject to the foregoing General and Specific Objections, OpenAI will make available for inspection, pursuant to an inspection protocol to be negotiated between the parties, the pretraining data for models used for ChatGPT that it locates after a reasonable search.

REQUEST FOR PRODUCTION NO. 15:

Documents concerning any commercial or supercomputing collaborations between and among the Defendants relating to the Generative AI Models, including licensing, developing, training, hosting, engineering, testing, deploying, or operating Generative AI Product(s) and Service(s).

RESPONSE TO REQUEST FOR PRODUCTION NO. 15:

OpenAI objects to the production of confidential and sensitive information before a protective order has been entered in this case. OpenAI objects to the production of electronically stored information before an applicable ESI protocol has been entered. OpenAI objects to the

phrases “commercial or supercomputing collaborations” as vague and ambiguous. OpenAI objects to this as overly broad, unduly burdensome, disproportionate to the needs of the case to the extent this request seeks documents regarding “commercial or supercomputing collaborations between and among the Defendants” unrelated to ChatGPT. OpenAI further objects to this request as seeking information and/or documents that are not relevant to any claim or defense in this litigation to the extent this request seeks documents regarding models not used for ChatGPT and/or that are not proportional to the needs of this case.

Subject to the foregoing General and Specific Objections, OpenAI states that it is willing to meet and confer regarding an appropriately narrow scope for this request.

Dated: San Francisco, California
March 25, 2024

MORRISON & FOERSTER LLP

By: /s/ Joseph C. Gratz
Joseph C. Gratz (*pro hac vice*)
JGratz@mofo.com
Vera Ranieri (*pro hac vice*)
VRanieri@mofo.com
425 Market Street
San Francisco, CA 94105-2482
Telephone: 415.268.7000
Facsimile: 415.268.7522

Allyson R. Bennett (*pro hac vice*)
ABennett@mofo.com
Rose S. Lee (*pro hac vice*)
RoseLee@mofo.com
MORRISON & FOERSTER LLP
707 Wilshire Boulevard
Los Angeles, CA 90017-3543
Telephone: 213.892.5200
Facsimile: 213.892.5454

Attorneys for Defendants
OPENAI, INC., OPENAI LP, OPENAI
GP, LLC, OPENAI, LLC, OPENAI
OPCO LLC, OPENAI GLOBAL LLC,
OAI CORPORATION, LLC, and
OPENAI HOLDINGS, LLC

CERTIFICATE OF SERVICE

I, the undersigned, hereby certify that on March 24, 2024, a true and correct copy of the foregoing document was served via electronic mail upon:

Genevieve Vose Wallace Ian B. Crosby Katherine Marie Peaslee SUSMAN GODFREY LLP 401 Union Street, Suite 3000 Seattle, WA 98101 Phone: 206-516-3880	Davida Brook Ellie Rae Dupler Emily K. Cronin SUSMAN GODFREY LLP 1900 Avenue of the Stars, Suite 1400 Los Angeles, CA 90067 Phone: 310.798.3150
Alexander Frawley Eudokia Spanos Tamar Lusztig Elisha Brandis Barron SUSMAN GODFREY LLP 1301 Avenue of the Americas, 32 nd Floor New York, NY 10019 Phone: 212.336.8330	Jennifer Maisel Kristen Logan Steven M Lieberman ROTHWELL FIGG ERNST & MANBECK, P.C. 901 New York Avenue, N.W.; Suite 900 East Washington, DC 20001 Phone: 202.783.6040
NYT-AI-SG-Service@simplelists.susmangodfrey.com	
<i>Attorneys for New York Times Company</i>	

Annette L. Hurst ORRICK, HERRINGTON & SUTCLIFF LLP The Orrick Building 405 Howard Street San Francisco, CA 94105 Phone: 415.773.5700	Christopher J. Cariello ORRICK, HERRINGTON & SUTCLIFF LLP 51 West 2nd Street New York, NY 10019 Phone: 212.506.5000
Jeffrey S. Jacobson FAEGRE DRINKER BIDDLE & REATH 1177 Avenue of the Americas New York, NY 10036 Phone: 212.248.3191 Email: jeffrey.jacobson@faegredrinker.com	Jared B. Briant FAEGRE DRINKER BIDDLE & REATH 1144 15 th Street, Suite 3400 Denver, CO 80202 Phone: 303.607.3588 Email: jared.briant@faegredrinker.com
NewYorkTimes_Microsoft_OHS@orrick.com;	
<i>Attorneys for Defendant Microsoft Corporation</i>	

By: /s/ heathersmith
Heather Smith